

R. Mitra

Re-run

RLS entered Page 1 of 8
Em 4/24/03



RAW SEQUENCE LISTING

DATE: 04/24/2003

PATENT APPLICATION: US/09/105,117K

TIME: 17:19:11

Input Set : N:\paola\US09105117K.raw.txt

Output Set: N:\CRF4\04242003\I105117K.raw

C--> 1 <110> APPLICANT: Forschungszentrum Juelich GmbH; Marina Vrlijc et al.
2 <120> TITLE OF INVENTION: Process for the microbial production of amino acids by
3 boosted activity of export carriers
4 <130> FILE REFERENCE: 1
5 <140> CURRENT APPLICATION NUMBER: US/09/105,117K
6 <141> CURRENT FILING DATE: 1998-06-17
7 <150> PRIOR APPLICATION NUMBER: PCT/DE96/02485
8 <151> PRIOR FILING DATE: 1996-12-18
9 <160> NUMBER OF SEQ ID NOS: 5
10 <170> SOFTWARE: PatentIn Ver. 2.0
11 <210> SEQ ID NO: 1
12 <211> LENGTH: 2374
13 <212> TYPE: DNA
14 <213> ORGANISM: Corynebacterium glutamicum
15 <220> FEATURE:
16 <221> NAME/KEY: gene
17 <222> LOCATION: CDS (1016)..(1726)
18 <223> OTHER INFORMATION: (LysE)
19 <400> SEQUENCE: 1
20
21 ccatttgctg aaggtgttac tctgcctggc ccaatttcctg cgggcgaaga agtgaaaaac 60
22 cctgaacctt ttcagaagta actaaggccg caatccctcg attgctgcat caacgacggc 120
23 gtctgtgagt ctagctagag atctagattc caggcgccat cgttgccaat acatcggtgt 180
24 gtcaatgggt atctcatcga ggaggatcac ttctcctgct tttagcatgg gagcagcttg 240
25 ggtttcggga agaagtcccc aaccaaggcc tcggcgaatt gcctcaccaa aaccttccgc 300
26 cgacgggaca atggatacgc gcctgcgccc cacaggacca tcgacgcgcc cgtccaggtc 360
27 acggtcttga agcacatctt tgggaccgaa gcgtaagacg ggcatcgcat cccaatctag 420
28 tttcccatca accatgtagg catcccgcac cacgggatta gcaatggcca agtggcgcat 480
29 ggttccaagt tctactactt cacatcccgc cacgggatta gcttcacggg ttaccgctcc 540
30 taaaacatct ccacgcccga gcaaggataa tgtgtgcgct tcatcttcca agcgcagcgt 600
31 gagcgttgct ccaccccaag aagctacctc gttgaacacg ggaggaaacc atgtggatag 660
32 cgaatctgcg ttgatggcga tggttaacgg gatttcagca aggcgtccag atagttgcgc 720
33 tttagtttct gcttgacgca acaccatttt ccgcgctgct tgcacaagga cttcaccgcg 780
34 ttcggttgct ttggccggtt ggggtgcgca taccaacact cgacccacgt gatgctcgag 840
35 agctttaacg cgctgactca ccgcccaggg ggaaatggaa agggctaagg aggcgccttc 900
36 gaagctgcct tcatcaatga ttgagagcaa agtgtccagt tgaatggggt tcatgaagct 960
37 atattaaacc atgttaagaa ccaatcattt tacttaagta cttccatagg tcacg atg 1018
38 Met
39 1
40 gtg atc atg gaa atc ttc att aca ggt ctg ctt ttg ggg gcc agt ctt 1066
41 Val Ile Met Glu Ile Phe Ile Thr Gly Leu Leu Leu Gly Ala Ser Leu
42 5 10 15
43 tta ctg tcc atc gga ccg cag aat gta ctg gtg att aaa caa gga att 1114
44 Leu Leu Ser Ile Gly Pro Gln Asn Val Leu Val Ile Lys Gln Gly Ile

ENTERED

RAW SEQUENCE LISTING

DATE: 04/24/2003

PATENT APPLICATION: US/09/105,117K

TIME: 17:19:11

Input Set : N:\paola\US09105117K.raw.txt

Output Set: N:\CRF4\04242003\I105117K.raw

```

45          20          25          30
46 aag cgc gaa gga ctc att gcg gtt ctt ctc gtg tgt tta att tct gac 1162
47 Lys Arg Glu Gly Leu Ile Ala Val Leu Leu Val Cys Leu Ile Ser Asp
48          35          40          45
49 gtc ttt ttg ttc atc gcc ggc acc ttg ggc gtt gat ctt ttg tcc aat 1210
50 Val Phe Leu Phe Ile Ala Gly Thr Leu Gly Val Asp Leu Leu Ser Asn
51          50          55          60          65
52 gcc gcg ccg atc gtg ctc gat att atg cgc tgg ggt ggc atc gct tac 1258
53 Ala Ala Pro Ile Val Leu Asp Ile Met Arg Trp Gly Gly Ile Ala Tyr
54          70          75          80
55 ctg tta tgg ttt gcc gtc atg gca gcg aaa gac gcc atg aca aac aag 1306
56 Leu Leu Trp Phe Ala Val Met Ala Ala Lys Asp Ala Met Thr Asn Lys
57          85          90          95
58 gtg gaa gcg cca cag atc att gaa gaa aca gaa cca acc gtg ccc gat 1354
59 Val Glu Ala Pro Gln Ile Ile Glu Glu Thr Glu Pro Thr Val Pro Asp
60          100          105          110
61 gac acg cct ttg ggc ggt tgc gcg gtg gcc act gac acg cgc aac cgg 1402
62 Asp Thr Pro Leu Gly Gly Ser Ala Val Ala Thr Asp Thr Arg Asn Arg
63          115          120          125
64 gtg cgg gtg gag gtg agc gtc gat aag cag cgg gtt tgg gta aag ccc 1450
65 Val Arg Val Glu Val Ser Val Asp Lys Gln Arg Val Trp Val Lys Pro
66          130          135          140          145
67 atg ttg atg gca atc gtg ctg acc tgg ttg aac ccg aat gcg tat ttg 1498
68 Met Leu Met Ala Ile Val Leu Thr Trp Leu Asn Pro Asn Ala Tyr Leu
69          150          155          160
70 gac gcg ttt gtg ttt atc ggc ggc gtc ggc gcg caa tac ggc gac acc 1546
71 Asp Ala Phe Val Phe Ile Gly Gly Val Gly Ala Gln Tyr Gly Asp Thr
72          165          170          175
73 gga cgg tgg att ttc gcc gct ggc gcg ttc gcg gca agc ctg atc tgg 1594
74 Gly Arg Trp Ile Phe Ala Ala Gly Ala Phe Ala Ala Ser Leu Ile Trp
75          180          185          190
76 ttc ccg ctg gtg ggt ttc ggc gca gca gca ttg tca cgc ccg ctg tcc 1642
77 Phe Pro Leu Val Gly Phe Gly Ala Ala Ala Leu Ser Arg Pro Leu Ser
78          195          200          205
79 agc ccc aag gtg tgg cgc tgg atc aac gtc gtc gtg gca gtt gtg atg 1690
80 Ser Pro Lys Val Trp Arg Trp Ile Asn Val Val Val Ala Val Val Met
81          210          215          220          225
82 acc gca ttg gcc atc aaa ctg atg ttg atg ggt tag ttttcgcggg 1736
83 Thr Ala Leu Ala Ile Lys Leu Met Leu Met Gly
84          230          235
85 ttttggaatc ggtggccttc gcccaaatgt tgatgccggc gtcgtgggaa atctcatcga 1796
86 tcgcctccaa ctcggcgctca gaaaactcca agttgttgag tgaatcaagg ctgttgtcca 1856
87 gctgctcaac tgacgaagca ccaatcaatg cactgggtcac ggatatccgcg ccgtactctc 1916
88 cttgctcgcg cagcaccat gcaagcgcca tctgcgcaag tgactgcccg cggttcctggg 1976
89 cgatgtcatt gagcttgccg accatatcaa tattgttcac gttcaacatg ccctcagaca 2036
90 gggacttacc ctggctggcg cgggaaccct ctggaattcc atcgagatat ttgtccgtga 2096
91 gcaggccctg cgcaagtggg gagaaagcaa tgacgccaag accattgttg gcagctgact 2156
92 gcaacaagtt ctcaccgtca tcgcccgggt cctccacca acgattaatg atggaatagc 2216
93 ttggctgatg aatcagaagc gggcagccct cctccgccat gaactcagcc gcctccgctg 2276

```

RAW SEQUENCE LISTING

DATE: 04/24/2003

PATENT APPLICATION: US/09/105,117K

TIME: 17:19:11

Input Set : N:\paola\US09105117K.raw.txt

Output Set: N:\CRF4\04242003\I105117K.raw

```

94      tgagctctgg accgtaggaa gaaatacccca cgtaaagagc ctttccagac gcaacaatgt 2336
95      cacgcaatgc gtacatgggt tcttccaaag gagtatct 2374
97 <210> SEQ ID NO: 2
98 <211> LENGTH: 236
99 <212> TYPE: PRT
100 <213> ORGANISM: Corynebacterium glutamicum
101 <220> FEATURE:
102 <223> OTHER INFORMATION: (LysE)
103 <400> SEQUENCE: 2
104      Met Val Ile Met Glu Ile Phe Ile Thr Gly Leu Leu Leu Gly Ala Ser
105           1             5             10             15
106      Leu Leu Leu Ser Ile Gly Pro Gln Asn Val Leu Val Ile Lys Gln Gly
107           20             25             30
108      Ile Lys Arg Glu Gly Leu Ile Ala Val Leu Leu Val Cys Leu Ile Ser
109           35             40             45
110      Asp Val Phe Leu Phe Ile Ala Gly Thr Leu Gly Val Asp Leu Leu Ser
111           50             55             60
112      Asn Ala Ala Pro Ile Val Leu Asp Ile Met Arg Trp Gly Gly Ile Ala
113           65             70             75             80
114      Tyr Leu Leu Trp Phe Ala Val Met Ala Ala Lys Asp Ala Met Thr Asn
115           85             90             95
116      Lys Val Glu Ala Pro Gln Ile Ile Glu Glu Thr Glu Pro Thr Val Pro
117           100            105            110
118      Asp Asp Thr Pro Leu Gly Gly Ser Ala Val Ala Thr Asp Thr Arg Asn
119           115            120            125
120      Arg Val Arg Val Glu Val Ser Val Asp Lys Gln Arg Val Trp Val Lys
121           130            135            140
122      Pro Met Leu Met Ala Ile Val Leu Thr Trp Leu Asn Pro Asn Ala Tyr
123           145            150            155            160
124      Leu Asp Ala Phe Val Phe Ile Gly Gly Val Gly Ala Gln Tyr Gly Asp
125           165            170            175
126      Thr Gly Arg Trp Ile Phe Ala Ala Gly Ala Phe Ala Ala Ser Leu Ile
127           180            185            190
128      Trp Phe Pro Leu Val Gly Phe Gly Ala Ala Ala Leu Ser Arg Pro Leu
129           195            200            205
130      Ser Ser Pro Lys Val Trp Arg Trp Ile Asn Val Val Val Ala Val Val
131           210            215            220
132      Met Thr Ala Leu Ala Ile Lys Leu Met Leu Met Gly
133           225            230            235
135 <210> SEQ ID NO: 3
136 <211> LENGTH: 2374
137 <212> TYPE: DNA
138 <213> ORGANISM: Corynebacterium glutamicum
139 <220> FEATURE:
140 <221> NAME/KEY: unsure
141 <222> LOCATION: CDS (2)..(652)
142 <223> OTHER INFORMATION: orf3
143      (complement to SEQ ID No. 1)
144 <220> FEATURE:

```

RAW SEQUENCE LISTING

DATE: 04/24/2003

PATENT APPLICATION: US/09/105,117K

TIME: 17:19:11

Input Set : N:\paola\US09105117K.raw.txt

Output Set: N:\CRF4\04242003\I105117K.raw

```

145 <221> NAME/KEY: gene
146 <222> LOCATION: CDS (1421)..(2293)
147 <223> OTHER INFORMATION: LysG
148 <400> SEQUENCE: 3
149     a gat act cct ttg gaa gaa acc atg tac gca ttg cgt gac att gtt gcg 49
150     Asp Thr Pro Leu Glu Glu Thr Met Tyr Ala Leu Arg Asp Ile Val Ala
151     1           5           10           15
152     tct gga aag gct ctt tac gtg ggt att tct tcc tac ggt cca gag ctc 97
153     Ser Gly Lys Ala Leu Tyr Val Gly Ile Ser Ser Tyr Gly Pro Glu Leu
154     20           25           30
155     aca gcg gag gcg gct gag ttc atg gcg gag gag ggc tgc ccg ctt ctg 145
156     Thr Ala Glu Ala Ala Glu Phe Met Ala Glu Glu Gly Cys Pro Leu Leu
157     35           40           45
158     att cat cag cca agc tat tcc atc att aat cgt tgg gtg gag gaa ccg 193
159     Ile His Gln Pro Ser Tyr Ser Ile Ile Asn Arg Trp Val Glu Glu Pro
160     50           55           60
161     ggc gat gac ggt gag aac ttg ttg cag tca gct gcc aac aat ggt ctt 241
162     Gly Asp Asp Gly Glu Asn Leu Leu Gln Ser Ala Ala Asn Asn Gly Leu
163     65           70           75           80
164     ggc gtc att gct ttc tca cca ctt gcg cag ggc ctg ctc acg gac aaa 289
165     Gly Val Ile Ala Phe Ser Pro Leu Ala Gln Gly Leu Leu Thr Asp Lys
166     85           90           95
167     tat ctc gat gga att cca gag ggt tcc cgc gcc agc cag ggt aag tcc 337
168     Tyr Leu Asp Gly Ile Pro Glu Gly Ser Arg Ala Ser Gln Gly Lys Ser
169     100          105          110
170     ctg tct gag ggc atg ttg aac gtg aac aat att gat atg gtc cgc aag 385
171     Leu Ser Glu Gly Met Leu Asn Val Asn Asn Ile Asp Met Val Arg Lys
172     115          120          125
173     ctc aat gac atc gcc cag gaa cgc ggg cag tca ctt gcg cag atg gcg 433
174     Leu Asn Asp Ile Ala Gln Glu Arg Gly Gln Ser Leu Ala Gln Met Ala
175     130          135          140
176     ctt gca tgg gtg ctg cgc gag caa gga gag tac ggc gcg gat acc gtg 481
177     Leu Ala Trp Val Leu Arg Glu Gln Gly Glu Tyr Gly Ala Asp Thr Val
178     145          150          155          160
179     acc agt gca ttg att ggt gct tcg tca gtt gag cag ctg gac aac agc 529
180     Thr Ser Ala Leu Ile Gly Ala Ser Ser Val Glu Gln Leu Asp Asn Ser
181     165          170          175
182     ctt gat tca ctc aac aac ttg gag ttt tct gac gcc gag ttg gag gcg 577
183     Leu Asp Ser Leu Asn Asn Leu Glu Phe Ser Asp Ala Glu Leu Glu Ala
184     180          185          190
185     atc gat gag att tcc cac gac gcc ggc atc aac att tgg gcg aag gcc 625
186     Ile Asp Glu Ile Ser His Asp Ala Gly Ile Asn Ile Trp Ala Lys Ala
187     195          200          205
188     acc gat tcc aaa acc cgc gaa aac taa cccatcaaca tcagtttgat 672
189     Thr Asp Ser Lys Thr Arg Glu Asn
190     210          215
191     ggccaatgcg gtcatacaca ctgccacgac gacgttgatc cagcgccaca ccttggggct 732
192     ggacagcggg cgtgacaatg ctgctgcgcc gaaacccacc agcgggaacc agatcaggct 792
193     tgccgcgaac gcgccagcgg cgaaaatcca ccgtccggtg tcgccgtatt gcgcgccgac 852

```

RAW SEQUENCE LISTING

DATE: 04/24/2003

PATENT APPLICATION: US/09/105,117K

TIME: 17:19:11

Input Set : N:\paola\US09105117K.raw.txt

Output Set: N:\CRF4\04242003\I105117K.raw

```

194 gccgccgata aacacaaacg cgtccaaata cgcattcggg ttcaaccagg tcagcacgat 912
195 tgccatcaac atgggcttta cccaaacccg ctgcttatcg acgtcacct ccacccgcac 972
196 ccggttgccg gtgtcagtg ccaccgccga accgccc aaa ggcgtgtcat cgggcacgg 1032
197 tggttctgtt tcttcaatga tctgtggcgc ttccaccttg tttgtcatgg cgtctttcgc 1092
198 tgccatgacg gcaaaccata acaggtaagc gatgccacc cagcgcataa tatcgagcac 1152
199 gatcggcgcg gcattggaca aaagatcaac gcccaagggt cggcgcatga acaaaaagac 1212
200 gtcagaaatt aaacacacga gaagaaccgc aatgagtcct tcgcgcttaa ttccttggtt 1272
201 aatcaccagt acattctgcg gtccgatgga cagtaaaaga ctggcccca aaagcagacc 1332
202 tgtaatgaag atttccatga tcaccatcgt gacctatgga agtacttaag taaaatgatt 1392
203 ggttcttaac atgggttaat atagcttc atg aac ccc att caa ctg gac act 1444
204                                     Met Asn Pro Ile Gln Leu Asp Thr
W--> 205                                     220                                     225
206 ttg ctc tca atc att gat gaa ggc agc ttc gaa ggc gcc tcc tta gcc 1492
207 Leu Leu Ser Ile Ile Asp Glu Gly Ser Phe Glu Gly Ala Ser Leu Ala
W--> 208                                     230                                     235                                     240
209 ctt tcc att tcc ccc tcg gcg gtg agt cag cgc gtt aaa gct ctc gag 1540
210 Leu Ser Ile Ser Pro Ser Ala Val Ser Gln Arg Val Lys Ala Leu Glu
W--> 211                                     245                                     250                                     255
212 cat cac gtg ggt cga gtg ttg gta tcg cgc acc caa ccg gcc aaa gca 1588
213 His His Val Gly Arg Val Leu Val Ser Arg Thr Gln Pro Ala Lys Ala
W--> 214                                     260                                     265                                     270
215 acc gaa gcg ggt gaa gtc ctt gtg caa gca gcg cgg aaa atg gtg ttg 1636
216 Thr Glu Ala Gly Glu Val Leu Val Gln Ala Ala Arg Lys Met Val Leu
W--> 217                                     275                                     280                                     285
218 ctg caa gca gaa act aaa gcg caa cta tct gga cgc ctt gct gaa atc 1684
219 Leu Gln Ala Glu Thr Lys Ala Gln Leu Ser Gly Arg Leu Ala Glu Ile
W--> 220                                     290                                     295                                     300                                     305
221 ccg tta acc atc gcc atc aac gca gat tcg cta tcc aca tgg ttt cct 1732
222 Pro Leu Thr Ile Ala Ile Asn Ala Asp Ser Leu Ser Thr Trp Phe Pro
W--> 223                                     310                                     315                                     320
224 ccc gtg ttc aac gag gta gct tct tgg ggt gga gca acg ctc acg ctg 1780
225 Pro Val Phe Asn Glu Val Ala Ser Trp Gly Gly Ala Thr Leu Thr Leu
W--> 226                                     325                                     330                                     335
227 cgc ttg gaa gat gaa gcg cac aca tta tcc ttg ctg cgg cgt gga gat 1828
228 Arg Leu Glu Asp Glu Ala His Thr Leu Ser Leu Leu Arg Arg Gly Asp
W--> 229                                     340                                     345                                     350
230 gtt tta gga gcg gta acc cgt gaa gct aat ccc gtg gcg gga tgt gaa 1876
231 Val Leu Gly Ala Val Thr Arg Glu Ala Asn Pro Val Ala Gly Cys Glu
W--> 232                                     355                                     360                                     365
233 gta gta gaa ctt gga acc atg cgc cac ttg gcc att gca acc ccc tca 1924
234 Val Val Glu Leu Gly Thr Met Arg His Leu Ala Ile Ala Thr Pro Ser
W--> 235                                     370                                     375                                     380                                     385
236 ttg cgg gat gcc tac atg gtt gat ggg aaa cta gat tgg gct gcg atg 1972
237 Leu Arg Asp Ala Tyr Met Val Asp Gly Lys Leu Asp Trp Ala Ala Met
W--> 238                                     390                                     395                                     400
239 ccc gtc tta cgc ttc ggt ccc aaa gat gtg ctt caa gac cgt gac ctg 2020
240 Pro Val Leu Arg Phe Gly Pro Lys Asp Val Leu Gln Asp Arg Asp Leu
W--> 241                                     405                                     410                                     415
242 gac ggg cgc gtc gat ggt cct gtg ggg cgc agg cgc gta tcc att gtc 2068

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 04/24/2003
PATENT APPLICATION: US/09/105,117K TIME: 17:19:12

Input Set : N:\paola\US09105117K.raw.txt
Output Set: N:\CRF4\04242003\I105117K.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 2

VERIFICATION SUMMARY

DATE: 04/24/2003

PATENT APPLICATION: US/09/105,117K

TIME: 17:19:12

Input Set : N:\paola\US09105117K.raw.txt

Output Set: N:\CRF4\04242003\I105117K.raw

L:5 M:270 C: Current Application Number differs, Wrong Format
L:205 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:208 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:211 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:214 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:217 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:220 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:223 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:226 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:229 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:232 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:235 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:238 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:241 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:244 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:247 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:250 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:253 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:256 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:259 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3